

差分量子化を用いた画像認識に関する研究

著者	李 菲菲
号	51
学位授与番号	3781
URL	http://hdl.handle.net/10097/37449

氏 名	李 菲 菲
授 与 学 位	博士（工学）
学 位 授 与 年 月 日	平成 19 年 3 月 27 日
学位授与の根拠法規	学位規則第 4 条第 1 項
研究科，専攻の名称	東北大学大学院工学研究科（博士課程）電子工学専攻
学 位 論 文 題 目	差分量子化を用いた画像認識に関する研究
指 導 教 官	東北大学教授 高橋 研
論 文 審 査 委 員	主査 東北大学教授 高橋 研 東北大学教授 川又 政征 東北大学教授 須川 成利 客員教授 大見 忠弘 東北大学助教授 小谷 光司 (未来科学技術共同研究センター)

論 文 内 容 要 旨

ネットワーク情報化社会の急速な発展により，電子取引，在宅勤務など，個々人の活動が多様化するため，個人認証技術が一層重要になってきている。認証技術には様々なものがあるが，顔，音声，指紋，虹彩などの身体特徴を用いる認証技術のなかでは，顔認識が最も自然で有効な認識手法である。従来の顔認識手法は，いずれも認識率が低い，処理が複雑，認識スピードが遅いといった欠点があった。そこで，著者は，「差分量子化を用いた画像認識」という新しい画像処理手法を提案した。これは，差分量子化により，画像情報から差分ベクトルの頻度分布という低次元の特徴量を抽出して用いる画像認識手法である。差分量子化を用いた顔認識処理のメリットとしては，

- (1) パターンの変動(平行移動)に不変な特徴である。
- (2) 簡単な計算式，少ない計算量によって顔認識処理が高速に実行できる。
- (3) 単独でも高い認識率であるが，従来の処理手法と異なる特徴量を用いた認識アルゴリズムであるため，従来の認識手法と組み合わせてさらなる認識能力の向上が可能である。

著者は，この原理に基づいた簡便かつ高い認識率を有する顔認識手法について研究を行うと共に，この手法をビデオシーン切り替えの自動検出手法に適用する研究を行った。本論文は，これらの研究成果をまとめたものであり，全文 5 章よりなる。

第 1 章は序論である。

第 2 章では，隣接画素差分量子化画像認識アルゴリズムについて述べている。すでに開発され，高い認識率を有しているベクトル量子化ヒストグラム法を用いた顔認識手法の本質を詳細に分析し，同等以上の機能をより簡便に実現する差分量子化画像認識手法を提案している。ベクトル量子化ヒストグラム法では，ベクトル量子化処理することにより，マッチング演算で 4x4 ブロック中の輝度変化方向と変化量を検出しているのである。これがベクトル量子化ヒストグラム法の本質である。同等機能は，隣接画素差分量子化を利用することにより，さらに簡単に実現できる。高い認識率を保ちながら，ビデオレートでの顔認識が可能となった。

本手法の概要としては，まず，図 1 に示すように，入力画像の各画素に対し，水平方向の隣接画素との輝度差 dI_x と垂直方向の隣接画素との輝度差 dI_y を差分演算により計算する。計算された dI_x ， dI_y

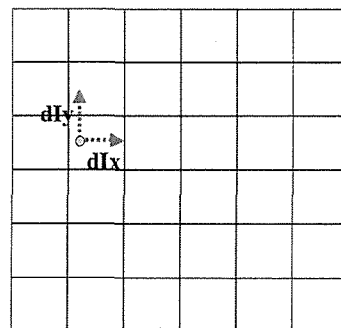


図 1.隣接画素差分計算

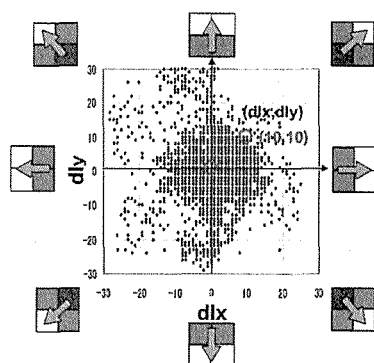


図 2. 輝度変化ベクトル分布

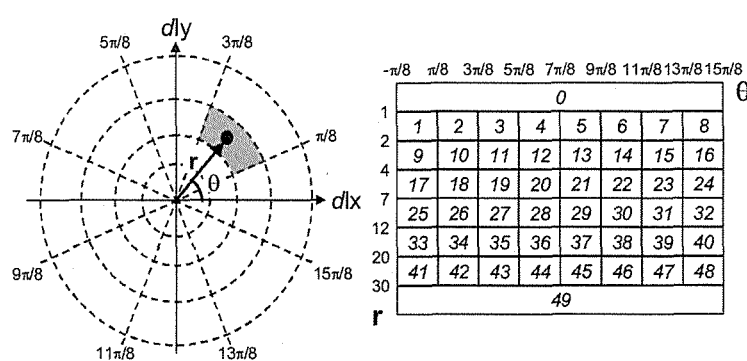


図 3. r - θ 空間での量子化テーブル

のペアは dIx - dIy 平面のひとつの輝度変化ベクトル点として表現できる。座標を直交座標から極座標に変換すると、角度 θ と距離 r は輝度変化の方向とその大きさを表すことになる。入力顔画像に含まれる全ての画素に対して輝度変化ベクトルを求め、 dIx - dIy 平面にプロットすると、差分ベクトル点の分布は、原点周辺に集中し、差分値が大きくなると、だんだん少なくなる傾向が見られる。これは人間顔面の輝度変化は大半が緩やかであることから説明できる。 dIx , dIy ベクトル点の分布の密度と形状が入力顔画像の特徴をよく表している(図 2)。次にそれぞれの輝度変化ベクトルを図 3 に示すような量子化テーブルを用いて、 r - θ 空間で量子化し、各量子化領域に含まれるベクトル点の数カウントする事で頻度分布を生成することにより特徴量抽出を行うものである。本章では、人物によりそれぞれ頻度分布が異なることを示し、この手法が顔認識に適用できることを明らかにしている。認識システムとして、登録の段階では、このヒストグラムが個人情報としてデータベースに登録される。認識の段階では、入力された顔画像よりヒストグラムが生成され、あらかじめ登録された個人毎のヒストグラムと比較され、最も距離の近いものが認識結果として出力される。マッチング指標としてはヒストグラム間のマンハッタン距離が用いられる。

引き続き第 3 章では、差分量子化を用いた顔認識の検証結果と認識精度をさらに向上させる手法の検討に関して論じている。まず様々な撮影条件を含む 40 人分の 400 枚の顔画像からなる小規模な公開の AT&T 顔画像データベースを用いて認識実験を行った。図 4 には認識結果を示している。横軸はフィルタサイズで、縦軸は認識率である。図に示すように、三本の曲線があるが、それぞれ 252 通り中の一番良いケース、悪いケース、と 252 通りの平均の値である。この図から分かるように、Low-Pass フィルタ処理により、認識率が向上になった。フィルタ 17×17 に 95.75% の最大平均認識率がある。これは VQ 顔認識率 95.6% と大体同じ認識率である。フィルタ処理は認識に不要なノイズ成分を除外することだけではなく、顔認識に重要な周波数成分を抽出するのにも非常に効率的であることが分かった。さらに認識結果を詳細に検証して認識手法の検討を行い、単一フィルタサイズで顔の周波数特徴情報の一部抽出できると考えれば、異なるフィルタサイズにより顔の異なる情報が抽出できると考えられる。複数のフィルタサイズの組み合わせでより多く有効な顔周波数特徴情報が抽出可能と考えられる。複数フィルタの組み合わせにより平均認識率を 98% に向上させた。量子化手法としてテーブル参

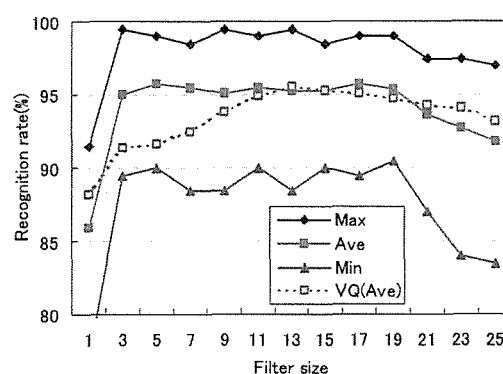


図 4. フィルタ処理による認識率の変化

照法を採用し、一枚の顔画像に対する全認識処理時間を、一般的なパーソナルコンピュータを用いた場合で 31 ms まで短縮した。

認識実験により、照明条件が認識に非常に大きな影響を与えることが分かった。光源方向の変化による影や明るさの影響を削減するために、線形関数当てはめにより近似的に照明条件を計算し、それと元画像の差分を取ることで、照明の影響を除外する手法を適用した。

光源方向の影響を除去するため、最も当てはまりの良い線形関数を計算し、近似の照明分布を計算する。その近似的な照明分布と画像との差分を取る。これにより、光源方向の変化により顔にできた影の影響を取り除いて、照明の影響を小さくする。明るさの影響を低減するために、差分画像内の平均と分散を一定とする。これで画像内の輝度値の分布が一定になる(図 5)。しかも、線形的な変換なので、元の有効な情報がそのまま残れる。照明の影響実環境下での認識実験において、図 6 に示すように、顔領域分割処理と複数フィルタの組み合わせにより、25 日までの時間経過に対して 100%の認識率を実現した。これは実用上極めて有益な成果である。

第 4 章では、差分量子化手法のビデオシーン切り替え検出への適用に関して論じている。差分量子化を用いることにより、高精度にビデオシーン切り替え検出が可能であることを示している。

差分量子化は以下の特徴を有する。一つは APIDQ

で生成された画像の頻度分布は画像の特徴を表す。もう一つは画面内の Object の位置が変わっても、頻度分布の変化が大きく変わらない。また、頻度分布の生成が速い(ビデオレート可能)。

ビデオは動画であるので、となりのフレーム同士は相関が高く、対応する頻度分布も近いはずである。シーンが変わると、頻度分布が大きく変化する。この差分量子化の特徴を利用して、ビデオの要約作成、シーンの検出などに応用できると考えられる。

ビデオシーン切り替え検出手順は図 7 のようになる。まず、入力ビデオ画像を用い、フィルタ処理を行う。これは移動平均フィルタを使う。それから、隣接画素差分計算をし、得られた差分値を極座標に変換し、その分布に対して量子化を行う。これにより、ヒストグラムが生成される。隣の画像フレーム同士頻度分布の類似度を計算して、ビデオシーン切り替え検出する。ヒストグラムの例から分かるように、となりのフレーム同士に対応する頻度分布は似ているが、シーンが変わると頻度分布が大きく変化する。

APIDQ で生成された頻度分布は画像の特徴を表している

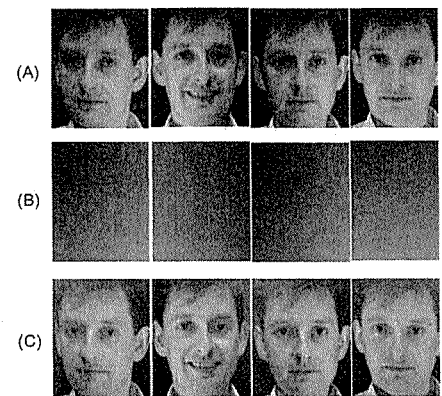


図 5. 照明条件の補正画像例

(A)元画像 (B)近似の照明分布
(C)平均と分散を一定とした画像

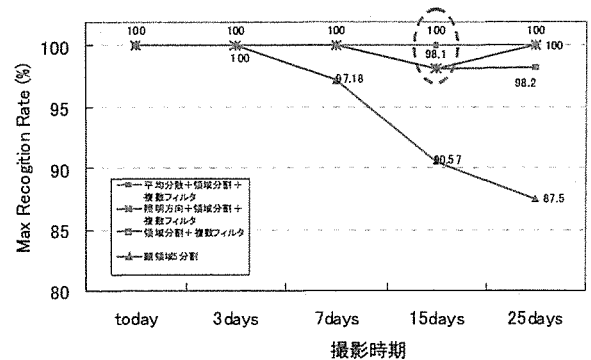


図 6. 平均と分散一定の認識結果

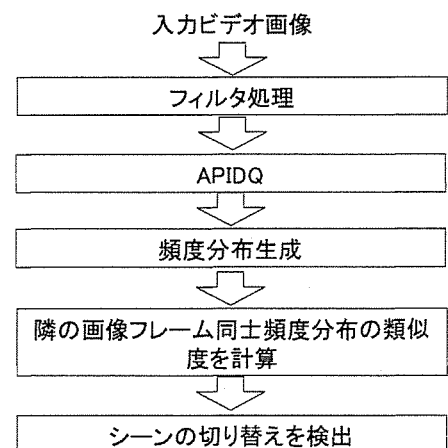


図 7. APIDQ を用いたシーン切り替え検出の手順

ことが分かる。APIDQ 法で計算したフレーム間の類似度の変化図からは、シーンが切り替えないフレーム間の類似度とシーン切り替えるフレーム間の類似度は離れていること分かる。フレーム間の類似度の変化からシーンの検出が可能である。また、APIDQ 法は従来の輝度ヒストグラム法により優れていることを示した。

さらに、動的しきい値設定手法を確立し、シーン切り替えの自動検出を実現した。シーン切り替えを検出するには、閾値を決めなければならない。ここまで検討した手法は、単一の固定閾値でシーン切り替えを検出できているが、フィルタのサイズ、画像サイズにより最適な一つの固定の閾値を選ぶのは難しい。それで、正規分布の性質を利用し、動的に適切な閾値を選ぶ手法について検討した。

また、現在標準的な動画圧縮フォーマットである MPEG データフォーマットの DC 画像に差分子量化を適用することで、完全な MPEG デコードを実行することなく圧縮データから直接シーン切り替え検出を行った。公開のビデオデータベースからいくつかの MPEG ビデオを用いて、従来手法と比べた。表 1 に示すように、フィルタ 5×5 を用いた場合の結果で、APIDQ 手法では誤検出された数と検出漏れ数がそれぞれ 2 カ所あるが、他の従来手法に比べ、かなり良い結果が得られた。直接 MPEG ビデオのシーン切り替え検出が可能であることを示した。これは、今後益々重要となる動画の自動分類技術実現に向けてその基礎となる重要な成果である。

第 5 章は結論である。

以上要するに本論文は、差分子量化画像認識技術を確立するとともに、高い認識率を有する顔認識手法、信頼性の高いビデオシーン切り替え自動検出手法を実現したものである。高度マルチメディア情報処理環境の構築に向けて意義の高いものであり、画像電子工学の発展に寄与するところが少なくない。

表 1. 各手法と APIDQ 法の比較結果

(T—正しく検出された数 F—誤検出された数
M—検出漏れ数)

	T	F	M
AIM	124	20	16
輝度ヒストグラム	122	30	18
LR	126	17	14
ME	130	10	10
APIDQ	138	2	2

論文審査結果の要旨

ネットワーク情報化社会の急速な発展により、電子取引、在宅勤務など、個々人の活動が多様化するため、個人認証技術が一層重要になってきている。認証技術には様々なものがあるが、顔、音声、指紋、虹彩などの身体特徴を用いる認証技術のなかでは、顔認識が最も自然で有効な認識手法である。従来の顔認識手法は、いずれも認識率が低い、処理が複雑、認識スピードが遅いといった欠点があった。そこで、著者は、「差分量子化を用いた画像認識」という新しい画像処理手法を提案した。これは、差分量子化により、画像情報から差分ベクトルの頻度分布という低次元の特徴量を抽出して用いる画像認識手法である。著者は、この原理に基づいた簡便かつ高い認識率を有する顔認識手法について研究を行うと共に、この手法をビデオシーン切り替えの自動検出手法に適用する研究を行った。本論文は、これらの研究成果をまとめたものであり、全文5章よりなる。

第1章は序論である。

第2章では、隣接画素差分量子化画像認識アルゴリズムについて述べている。すでに開発され、高い認識率を有しているベクトル量子化ヒストグラム法を用いた顔認識手法の本質を詳細に分析し、同等以上の機能をより簡便に実現する差分量子化画像認識手法を提案している。この手法は、入力顔画像に含まれる全ての画素に対して隣接画素差分演算により輝度変化ベクトルを求め、そのベクトル分布を量子化して頻度分布を生成することにより特徴量抽出を行うものである。人物によりそれぞれ頻度分布が異なることを示し、この手法が顔認識に適用できることを明らかにしている。

第3章では、差分量子化を用いた顔認識の検証結果と認識精度をさらに向上させる手法の検討に関して論じている。まず小規模な公開の AT&T 顔画像データベースを用いて認識実験を行い、適切なフィルタを用いることで、様々な撮影条件を含む40人分の400枚の顔画像に対し、95.7%の平均認識率を実現した。さらに認識結果を詳細に検証して認識手法の検討を行い、複数フィルタの組み合わせにより平均認識率を98%に向上させた。量子化手法としてテーブル参照法を採用し、一枚の顔画像に対する全認識処理時間を、一般的なパーソナルコンピュータを用いた場合で31msまで短縮した。さらに、線形関数近似による照明方向補正およびコントラスト補正により照明条件の影響を低減し、実環境下での認識実験において、顔領域分割処理と複数フィルタの組み合わせにより、25日までの時間経過に対して100%の認識率を実現した。これは実用上極めて有益な成果である。

第4章では、差分量子化手法のビデオシーン切り替え検出への適用に関して論じている。差分量子化を用いることにより、高精度にビデオシーン切り替え検出が可能であることを示している。さらに、動的しきい値設定手法を確立し、シーン切り替えの自動検出を実現した。また、現在標準的な動画圧縮フォーマットであるMPEGデータフォーマットのDC画像に差分量子化を適用することで、完全なMPEGデコードを実行することなく圧縮データから直接シーン切り替え検出が可能であることを示した。これは、今後益々重要となる動画の自動分類技術実現に向けてその基礎となる重要な成果である。

第5章は結論である。

以上要するに本論文は、差分量子化画像認識技術を確立するとともに、高い認識率を有する顔認識手法、信頼性の高いビデオシーン切り替え自動検出手法を実現したものである。高度マルチメディア情報処理環境の構築に向けて意義の高いものであり、画像電子工学の発展に寄与するところが少なくない。

よって、本論文は博士(工学)の学位論文として合格と認める。